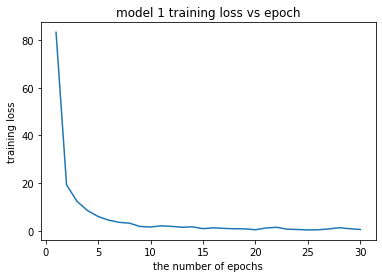
**Homework 2**

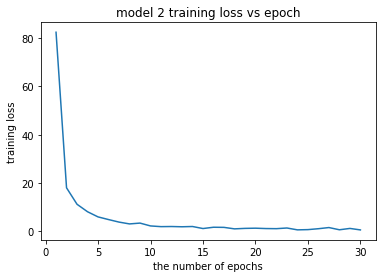
Due data: May 2, 2021

Draw a graph of training loss vs epoch (x-axis: the number of epochs, y-axis: training loss) for each model.

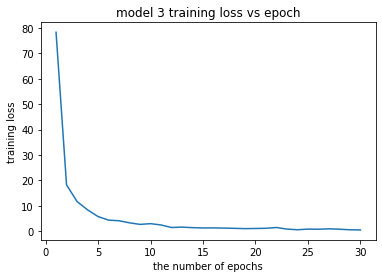
Model1)



Model2)

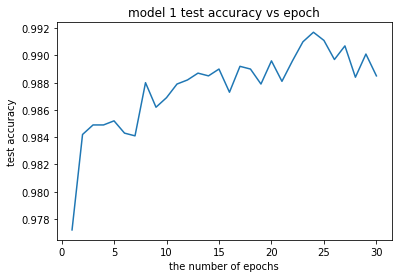


Model3)

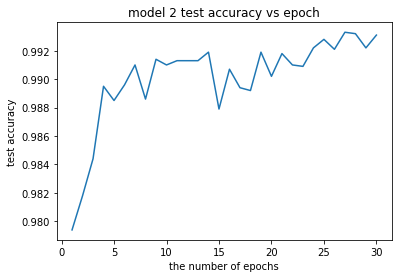


Draw a graph of test accuracy vs epoch (x-axis: the number of epochs, y-axis: test accuracy) for each model.

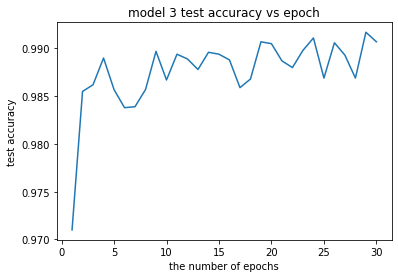
Model1)



Model2)



Model3)



Complete an evaluation table:

|  |  |  |
| --- | --- | --- |
|  | Final training loss | Final test accuracy |
| Model 1 | 0.505 | 98.9% |
| Model 2 | 0.529 | 99.3% |
| Model 3 | 0.556 | 99.1% |

Write a report including 3 training loss graphs, 3 test accuracy graphs, an evaluation table and an analysis & discussion (at least 10 lines).

In training loss graphs, all three graphs are drawing similar shape. Until training 5 epochs, training loss is rapidly decreased and after that, there is little change in training loss. While in test accuracy graphs, there are some variance between models. All 3 models increased accuracy to about 99%. However, graph of model1 shows the suggestion of overfitting after 25 epochs. Because it gets highest accuracy before 25 epochs and gradually decrease the accuracy after that. Graph of model2 shows quite rising line although quite irregular. In case of model3, the line of test accuracy graph goes back and forth at a certain level repeatedly.

Observing an evaluation table, Model2 seems most superior model although all models get similar and great test accuracy. I think that the reason why Model2 is best is related with the number of channels in convolution layers. Model2 has 128 channels at 4th convolution layer, which is much more than Model1 and Model3. I think if the more channels there are, the more detailed information can be contained in learning. So test accuracy can go slightly better.